

Simulation temporelle et spatiale des changements d'occupation du sol par modélisation stochastique

Jean François Mari¹, Arnaud Gobillot², Marc Benoît²

¹ Loria, Université de Lorraine, CNRS, Inria
Campus scientifique BP 239

F- 54506 Vandoeuvre-lès-Nancy Cedex, France

² INRA-SAD / Unité ASTER (Agro-Systèmes, Territoires)
662 avenue Louis Buffet F- 88500 Mirecourt, France

December 12, 2017

Abstract

The landscape patterns are the results of the human activity that adapts the land cover and its use (LUC) to constraints or economic opportunities encountered by various actors. The temporal and spatial arrangement of LUC has a strong influence on environmental risks. Its knowledge allows to develop prospective simulation models to assess the impact of different scenarios on the evolution of the land use and land cover change (LUCC). This paper presents a methodology based on stochastic models to identify, locate and simulate the temporal LUCC of regions differentiated by their successions by means of “neutral” models. The concept of a “neutral” model has been introduced in ecology for describing the characteristics of coexisting populations in a territory. Every individual - whatever its species it belongs to in the ecosystem - undergoes the same constraints without taking precedence over another. The result is a balance that is regulated by processes whose dynamics do not vary. In the framework of spatially explicit models coupled to a geographical information system according to the categorization defined by [AGG⁺00], the territory is described as an assembly of homogeneous discontinuous units - called patches - represented by a regular grid of pixels.

We propose a territorial simulation method that takes into account the dynamics of temporal and spatial patterns of previous successions observed in different sub-regions. We carry out a classification of the territory in homogeneous patches with respect to their LUC. The simulation of the LUC succession is then made on these patches at the pixel level taking into account the rotations / successions that were practiced the previous two years. The differences

between simulated and observed data make it possible to detect breaks in the process of developing the territory according to the constraints or opportunities for farmers. Differences are measured using the pixel recognition rate on a rasterized representation of the territory. We assess our method on data that come from surveys in the Vittel-Contrexéville territory (East of France) which is subject to water quality issues.

Résumé

L'organisation d'un territoire est le reflet de l'activité humaine qui y adapte la couverture du sol et son usage en fonction des contraintes ou opportunités économiques. L'arrangement temporel et spatial des occupations du sol (OCS) a une forte influence sur les risques environnementaux. Sa connaissance permet d'élaborer des modèles de simulation prospective afin d'évaluer l'impact de différents scénarios d'évolution de la couverture du sol. Cet article présente une méthodologie à base de modèles stochastiques pour identifier, localiser et simuler les occupations temporelles des régions différenciées par leurs successions de culture dans la perspective de construction d'un modèle "neutre". Le concept de modèle "neutre" a été introduit en écologie pour décrire les caractéristiques des populations coexistant dans un territoire. Les individus – quelque soit leur nature – endurent les mêmes contraintes sans qu'une espèce ne prenne le pas sur une autre. Il en résulte un équilibre qui est régulé par des processus dont la dynamique ne varie pas. Nous nous plaçons dans un cadre de définition de modèles spatialement explicites couplés à un système d'information géographique selon la catégorisation définie par Agarwal [AGG⁺00]. Le territoire est décrit comme un assemblage d'unités homogènes discontinues – appelées patches – lesquelles sont constituées d'une grille régulière de pixels. Nous effectuons préalablement une classification du territoire d'étude en patches homogènes vis à vis des successions de culture qui s'y pratiquent. La simulation de la succession est ensuite faite sur ces patches au niveau du pixel en tenant compte des rotations / successions qui se pratiquaient les deux années précédentes. Nous proposons ainsi une méthode de simulation du territoire qui tienne compte des dynamiques temporelles et spatiales des successions précédentes observées dans différentes sous régions. L'écart entre données simulées et observées permet de déceler des ruptures dans le processus de mise en valeur du territoire en fonction des contraintes ou opportunités s'offrant aux agriculteurs. Les écarts sont mesurés à l'aide du taux de reconnaissance de pixels sur une représentation rasterisée du territoire. Les données utilisées dans cette étude proviennent d'enquêtes effectuées sur le territoire de Vittel-Contrexéville concernant la qualité des eaux souterraines.

mots clés : simulation prospective, modèle neutre, HMM, LUCC, paysage, protection de l'eau

keywords : prospective simulation, neutral model, HMM, land use change, landscape, water protection

1 Introduction

L’organisation d’un territoire est le reflet de l’activité humaine qui y adapte la couverture du sol, leurs successions dans le temps et leurs usages en fonction des contraintes ou opportunités techniques et économiques [AGG⁺00]. Cette organisation se traduit sous la forme d’un paysage qui agrège ces diverses activités [BRM⁺12].

L’arrangement temporel et spatial des occupations du sol (OCS) a une forte influence sur les risques environnementaux [For95, JSMP06, SLOW11]. La fouille des données agronomiques a fait émerger des modèles de classification des mosaïques parcellaires et des nouveaux outils de visualisation des résultats [LMB09] qui permettent d’extraire des connaissances sur les changements d’usage et de couverture des sols. Comme dans tous processus vivants, la variabilité des données captées sur un territoire peut être cernée par des modèles stochastiques qui font apparaître des régularités explicables par un expert sous réserve d’une visualisation adéquate [CMA⁺10, SLM⁺12]. La connaissance extraite permet d’alimenter des modèles de simulation prospective afin d’évaluer l’impact des états actuels et de différents scénarios d’évolution de la couverture du sol [VL01, EDB12].

La définition d’un modèle de simulation prospective d’un territoire pose le problème de la description de l’organisation temporelle et spatiale du territoire [Hou15]. La calibration du modèle pose le problème du choix de la période temporelle et de la zone d’étude pour l’apprentissage des paramètres. Enfin sa validation amène sa comparaison à un modèle “neutre” qui représente la poursuite de la tendance actuelle ne serait ce que pour y déceler des ruptures dans la dynamique paysagère [PHD04].

Le concept de modèle “neutre” a été introduit en écologie par Hubell [Hub01] pour décrire les caractéristiques des populations coexistant dans un territoire. Les individus – quelque soit leur nature – endurent les mêmes contraintes sans qu’une espèce ne prenne le pas sur une autre. Il en résulte un équilibre qui est régulé par des processus dont la dynamique ne varie pas. Lorsque le modèle est essentiellement statistique, il est confondu avec le modèle “nul” ou d’ordre zéro qui représente l’hypothèse nulle H_0 . Pourtant, Gotelli [GM06] signale des différences dans la fonction de ces deux modèles. Dans notre étude, en Agronomie des territoires, comme présenté par Gaucherel [GFA⁺06], nous considérons que les populations sont les occupations du sol qui s’expriment dans des rotations et colonisent des parties – “patches” – du territoire.

Cet article présente plusieurs modèles neutres à base de modèles stochastiques pour identifier, localiser et simuler en fonction du temps les occupations des régions différenciées par leurs successions de culture. Nous nous plaçons dans un cadre de simulations inductives d’un territoire à l’aide de modèles spatialement explicites couplés à un système d’information géographique selon la catégorisation définie par Agarwal [AGG⁺00]. Le territoire est décrit comme un assemblage d’unités homogènes discontinues – appelées patches – lesquelles sont constituées d’un treillis régulier de pixels sous la forme d’une image *raster*. Nous effectuons préalablement une classification du territoire d’étude en patches

homogènes vis à vis des successions de culture qui s’y pratiquent. La simulation des successions de culture est ensuite faite dans ces patches au niveau du pixel. Nous proposons ainsi une méthode de simulation du territoire inductive qui part des dynamiques d’observations ponctuelles temporelles et spatiales des années précédentes observées dans différentes sous régions. L’écart entre données simulées et observées permet de déceler des ruptures dans le processus de mise en valeur du territoire en fonction des contraintes ou opportunités s’offrant aux agriculteurs. Les écarts sont mesurés à l’aide du taux de reconnaissance de pixels sur une représentation rasterisée du territoire.

Cet article est organisé de la façon suivante : après cette introduction, la section 2 décrit les données d’étude ainsi que les modèles stochastiques – modèles de Markov cachés d’ordre supérieur et champs de Markov – utilisés pour segmenter et classer la mosaïque des parcelles agricoles. Ces modèles sont implémentés dans les systèmes CarottAge et ARPEnTAge qui sont brièvement présentés car plus précisément décrits dans [LBBS⁺06, MLB13]. Cette section se termine par la présentation de notre méthode pour prolonger une séquence de rotation de cultures à partir de l’historique de la parcelle qui la porte. Enfin la section “résultats” montre comment l’utilisation d’ARPEnTAge sur les données d’enquête sur le territoire de Vittel-Contrexéville soumis à des risques de qualité des eaux souterraines permet de localiser des sous régions de même dynamique et donne des résultats de simulation plus contrastés en éliminant l’influence des zones stables (forêts, prairies permanentes) et permet de définir plus précisément un modèle “neutre” dans lequel siègent des rotations caractérisant la tendance actuelle. Différents modèles neutres de complexité croissante y sont comparés. Finalement, la section “discussion” situe notre travail dans la communauté de la simulation prospective en discutant son apport et ses limitations.

2 Matériel et méthodes

2.1 Les données relatives au territoire de Vittel-Contrexéville

Les données utilisées dans cette étude ont été obtenues dans le cadre du projet AGREV 3 (Agriculture Environment Vittel) qui accompagne Agrivair – filiale de *Nestle waters* – dans ses actions de protection des ressources naturelles d’eau minérale naturelle. Pour être ainsi nommées les eaux embouteillées ne peuvent pas contenir plus de 10 mg de nitrate par litre. Pour atteindre cet objectif sur un territoire de 6200 ha de terre agricole, il est nécessaire de mettre en place et d’évaluer des pratiques agricoles et donc, de dessiner un paysage porteur d’une agriculture viable qui permette une protection sur le long terme des ressources en eaux minérales naturelles. Ainsi, deux pressions agricoles sont à caractériser : l’assolement sur le temps long ainsi que les pratiques de fertilisation. Le rôle clé des successions dans la protection a pu être identifié sur le territoire depuis le début du programme de recherche en 1989. Dans cet article, nous verrons comment un modèle “neutre” détecte les modifications de successions d’occupations du sol en fonction d’une politique environnementale.

"non Classé"	X	0
"prairie permanente"	Mead	1
"tournesol"	SunF	2
"céréale"	Wheat	3
"colza"	Rape	4
"maïs"	Corn	5
"chaume"	Fall	6
"prairie temporaire"	t_Me	7
"vergers"	Orch	8
"bois"	Forest	9
"sorgho"	Sorgh	10
"luzerne"	Luze	11

Table 1: Les occupations du sol relevées par enquêtes de terrain et leurs abréviations sur les diagrammes

Le territoire de Vittel Contrexéville – 78 km^2 – est enquêté depuis 1996 et sa représentation parcellaire cartographiée dans plusieurs SIG. Au cours du temps, l’étendue du territoire enquêté ainsi que la classification des occupations des parcelles ont évolué. Certaines modalités sont apparues et d’autres abandonnées. Dans notre travail, le territoire est représenté par une grille régulière de pixels – ou sites – avec une résolution de 15 m. Chaque site possède une modalité parmi les 11 retenues (cf. Tab. 1). L’ensemble constitue un gros corpus d’informations temporelles et spatiales possédant un niveau de détails supérieur à ce qu’une analyse d’images satellites peut actuellement obtenir. Ce corpus est représenté sous la forme d’une matrice constituée d’autant de lignes qu’il y a de sites dans toutes les parcelles enquêtées, une ligne regroupant les différentes occupations temporelles de ce site.

2.2 Les logiciels CarottAge et ARPEnTAge

Nos efforts de modélisation de la mosaïque parcellaire porteuse de successions d’OCS à des fins d’extraction de connaissance se sont traduits par la réalisation de deux boîtes à outils logiciels : CarottAge et ARPEnTAge. Ces deux logiciels modélisent les processus temporels et spatiaux d’allocation des OCS sur les parcelles constituant un territoire en se fondant sur le même principe stochastique de localité. Pour simplifier la modélisation, nous acceptons l’hypothèse markovienne qui stipule que, sur une parcelle, la règle de succession ne dépend que de l’occupation actuelle de la parcelle et de l’occupation des deux années précédentes ainsi que des occupations des parcelles voisines. Ces hypothèses nous permettent d’utiliser les modèles de Markov d’ordre supérieur pour les séries ainsi que des champs aléatoires de Markov en dimension deux (MRF comme *Markov Random Field*) pour les images. Les experts agronomes acceptent facilement ces hypothèses. Plus l’ordre du modèle de Markov est grand, plus celui-ci peut modéliser des longues successions et plus il intéresse l’agronome. Mais

l’aspect intentionnel de la séquence qui concrétise le travail intelligent de mise en valeur du territoire n’est pas pris en compte.

Stochastique provient du grec *stokhastikos* : qui cerne bien et qualifie tout ce qui relève de la statistique et du calcul des probabilités. Dans une optique Bayésienne, la théorie des probabilités mesure une incertitude et permet de mieux cerner ce qui est dû au hasard de ce qui est explicable. Cette façon de procéder est complémentaire d’une approche analytique et explicative fondée sur un mécanisme de raisonnement. En mesurant précisément par une probabilité ce que l’on qualifie au premier abord d’aléatoire, on diminue l’indéterminisme de notre perception du processus et on peut faire apparaître des comportements explicables donc prévisibles qui pourront être réutilisés dans un mécanisme de raisonnement et de simulation. Ce mécanisme d’extraction et de réutilisation est un des principes d’extraction de connaissances par la fouille de données.

CarottAge est un ensemble de programmes sous licence libre Gnu qui met en œuvre des modèles de Markov cachés pour classer et représenter des séries temporelles d’OCS. CarottAge est un rétro acronyme formé à partir des mots carotte – qui se traduit par markov en russe – et âge. C’est aussi un mécanisme de datation. Ce logiciel a été utilisé principalement pour caractériser les successions de cultures du bassin de la Seine [MSB07] et de la plaine de Niort [LMB09]. Il permet l’apprentissage automatique des modèles stochastiques à partir de corpus de données ainsi que la visualisation des probabilités de successions majoritaires comme présentées dans les diagrammes de la figure 5.

ARPEntAge est un acronyme signifiant : Analyse de Régularités dans les Paysages : Environnement, Territoires, Agronomie. Ce logiciel est bâti au-dessus de CarottAge et met en œuvre des MRF pour représenter un champ aléatoire de séries temporelles. Il a été présenté dans EMS [MLB13] et a été testé sur des images de télédétection et des relevés d’enquêtes à différentes échelles. ARPEntAge délivre des “shapefiles” localisant des patches ainsi que leurs diagrammes de Markov représentant leurs successions pertinentes en fonction de leurs probabilités (cf. figure 4 et diagrammes 5 et 6). Les outils de visualisation développés avec ARPEntAge permettent ainsi à un système d’information géographique de représenter l’évolution temporelle des occupations – simples ou rotations – constituant les différents patches.

Les paragraphes suivants vont décrire les modèles stochastiques pour la classification et les algorithmes mis en œuvre pour la simulation de modèles neutres.

2.3 Notations et définitions

Processus stochastique : Soit un ensemble T , engendré par l’indice t , et une variable aléatoire Q_t , on appelle processus stochastique l’ensemble des variables aléatoires $(Q_t, t \in T)$. En général, T est dénombrable et représente le temps, le processus est alors dit *discret*.

Suite stochastique : Soit \mathcal{E} un ensemble fini appelé ensemble d’états. Entre les instants t et $t + 1$, le système passe aléatoirement de l’état u à l’état v . On définit la suite stochastique $(Q_1, Q_2, \dots, Q_t, \dots)$ à ensemble discret

d'états par la relation :

$$(Q_t = u) \Leftrightarrow \text{le système est dans l'état } u \text{ au temps } t$$

Chaîne de Markov : Une chaîne de Markov est une suite stochastique à ensemble discret d'états telle que :

Suite du premier ordre :

$$\begin{aligned} \text{Prob}(Q_t = w / Q_{t-1} = v, Q_{t-2} = u, Q_{t-3} = \dots) \\ = \text{Prob}(Q_t = w / Q_{t-1} = v) \end{aligned}$$

Suite du deuxième ordre :

$$\begin{aligned} \text{Prob}(Q_t = w / Q_{t-1} = v, Q_{t-2} = u, Q_{t-3} = \dots) \\ = \text{Prob}(Q_t = w / Q_{t-1} = v, Q_{t-2} = u) \end{aligned}$$

Propriétés Une chaîne est dite *homogène* si les probabilités de transition ne dépendent pas du temps.

Si l'ensemble \mathcal{E} est fini et possède N états, $\mathcal{E} = \{e_1, e_2 \dots e_N\}$, alors on dit que la chaîne est *finie*. On peut alors poser :

chaînes du premier ordre

$$a_{ij} = \text{Prob}(Q_t = e_j / Q_{t-1} = e_i), \quad \sum_j a_{ij} = 1$$

chaînes du second ordre

$$a_{ijk} = \text{Prob}(Q_t = e_k / Q_{t-1} = e_j, Q_{t-2} = e_i), \quad \sum_k a_{ijk} = 1$$

2.4 Modèles de Markov cachés pour représenter les séries temporelles d'occupations du sol

Les séries temporelles d'occupations des sols sont décidées par de nombreux acteurs utilisant le territoire : agriculteurs, forestiers, particuliers ... L'OCS au temps t sur un site se justifie autant par les OCS aux instants $t-1$, $t-2$, ... au même endroit que par celles de ses voisins spatiaux au temps t . Ceci en particulier car des sites voisins disposent de caractéristiques proches en termes de pentes, sols, distances aux villes et villages.

Les modèles de Markov cachés (HMM pour *Hidden Markov Model*) dérivent des chaînes de Markov. Une chaîne de Markov définit un seul processus stochastique : elle possède un ensemble d'états – par exemple les occupations successives d'une parcelle – directement observables. Dans un HMM, un état n'est pas uniquement associé à une observation mais à une densité de probabilité sur l'ensemble des observations. Ainsi, on considère que la répartition — c.-à-d. la densité ou

l'assolement — des OCS évolue selon un processus de Markov. La répartition à un pas de temps donné ne dépend que de la répartition aux pas précédents suivant l'ordre du modèle.

Un modèle de Markov caché est défini par la donnée de :

une chaîne de Markov, sur un ensemble $\mathcal{E} = \{e_1, e_2, \dots, e_N\}$. On appelle \mathcal{A} la matrice donnant les probabilités de transition entre états : $\mathcal{A} = (a_{ij})$ pour un modèle d'ordre 1 (HMM1), $\mathcal{A} = (a_{ijk})$ pour un modèle d'ordre 2 (HMM2).

un processus observable, représenté par les densités des observations qui sont faites sur les états e_i . On note $b_i(o) = P(O_t = o/e_i)$, $1 \leq t \leq T$ la probabilité d'observer o dans l'état e_i à l'instant t .

La matrice \mathcal{A} contient un ensemble de valeurs qui permettent de définir la topologie du graphe des transitions entre états : quelles sont les transitions autorisées ? aller simple ($a_{ij} > 0$, $a_{ji} = 0$), aller-retour ($a_{ij} > 0$, $a_{ji} > 0$), bouclage ($a_{ii} \neq 0$), etc.

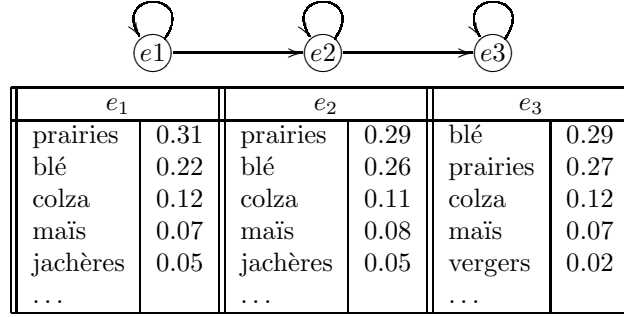


Figure 1: Un HMM représentant une suite d'assolements par les états cachés e_1, e_2 et e_3

En adoptant les notations définies dans [RJ95], on appelle $Q_1^T = q_1, q_2, \dots, q_T$ une suite d'états pris dans \mathcal{E} et atteints respectivement aux temps 1, 2, \dots T et $O_1^T = o_1, o_2, \dots, o_T$ une suite synchrone d'observations (la séquence d'OCS).

Dans le cas des modèles de Markov gauche-droite décrits dans la figure 1, l'alignement entre états et observations définit une segmentation temporelle en 3 périodes : $[1:t_2 - 1]$, $[t_2:t_3 - 1]$ et $[t_3:T]$.

t =	1	2	...	t_2	...	t_3	...	T
Processus caché =	Q_1	Q_2	...					Q_T
Suite d'états cachés =	e_1	e_1	...	e_1	e_2	...	e_2	e_3
Processus observé =	O_1	O_2	...					O_T
Suite d'OCS observées =	o_1	o_2	...					o_T

Figure 2: Relation entre suite d'états cachés et suite d'observations (les OCS)

Pour un HMM2, la probabilité de la suite d'états Q_1^T est définie dans l'équation

$$\text{Prob}(Q_1^T) = \Pi_{q_1} a_{q_1 q_2} \prod_{t=3}^T a_{q_{t-2} q_{t-1} q_t} \quad (1)$$

dans laquelle Π_{q_1} est la probabilité de l'état q_1 et $a_{q_1 q_2}$ la probabilité de la transition $q_1 \rightarrow q_2$ (initialisation du modèle aux instants $t = 1$ et $t = 2$).

La probabilité de l'alignement entre la suite d'états Q_1^T et la suite d'observations O_1^T est définie par la probabilité jointe $\text{Prob}(Q_1^T, O_1^T)$:

$$\text{Prob}(Q_1^T, O_1^T) = \Pi_{q_1} b_{q_1}(o_1) a_{q_1 q_2} b_{q_2}(o_2) \prod_{t=3}^T a_{q_{t-2} q_{t-1} q_t} b_{q_t}(o_t). \quad (2)$$

La vraisemblance $L(O)$ est calculée en considérant tous les alignements possibles.

$$L(O) = \sum_Q \text{Prob}(Q, O) \quad (3)$$

Comme le montre la figure 2, la modélisation d'une suite d'observations par des HMM est ainsi fondée sur deux principes : (i) la séquence peut être découpée en segments par une chaîne de Markov et (ii) chaque segment est la réalisation d'un processus stationnaire représenté par une densité de probabilité sur l'espace des observations. La modélisation par HMM souffre de deux inconvénients : (i) la longueur (ou la durée) d'un segment suit une loi géométrique. La probabilité d'un segment diminue avec sa longueur, ce qui n'est pas réaliste. (ii) La corrélation entre OCS successives n'est pas prise en compte. Nous aborderons les problèmes posés par ces limitations dans la partie discussion.

Une fois donné un corpus de données représentant les occupations du sol entre les instants 1 et T , différents algorithmes permettent l'apprentissage d'un HMM2 et la détermination des paramètres \mathcal{A} et $b_i(\cdot)$. Nous utilisons l'algorithme Forward - Backward [MHK97] qui est une variante de l'algorithme EM [DLR77]. Le résultat est constitué par les nouvelles valeurs des transitions a_{ijk} et des densités $b_i(\cdot)$.

Dans HMM tel que défini dans la figure 1, il n'est pas possible de mesurer la probabilité d'une succession de cultures puisqu'une culture n'apparaît qu'à l'intérieur d'une répartition constituant l'état. Pour pallier ce défaut, nous introduisons des états uniquement associés aux cultures majoritaires que nous voulons étudier (blé, maïs, orge, ...). Le HMM obtenu possède deux types d'états : les états "de réserve" qui sont associés normalement à des répartitions de cultures – comme dans le modèle Fig. 1 – et les états "de Dirac" qui sont associés à une culture seule et définis par une densité où la probabilité de cette culture vaut 1 et les probabilités de toutes les autres occupations valent 0. Finalement la figure 3 donne la topologie d'un modèle qui permet d'étudier les transitions entre les états associés au blé, orge, maïs, colza et prairie. Les états cachés sont dénotés 2, 3 et 4 et jouent le rôle d'états de réserve.

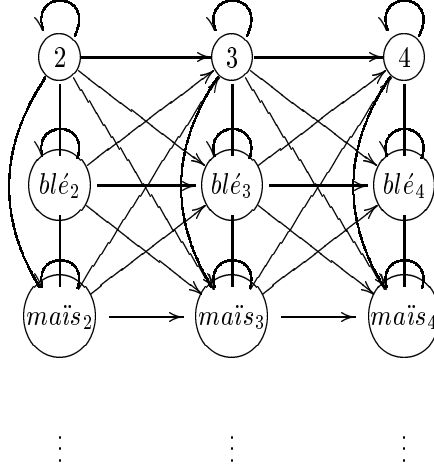


Figure 3: Les états notés 2, 3 et 4 sont associés à une distribution de cultures, contrairement aux états dénommés par une occupation. Le nombre de colonnes définit le nombre de périodes d’observation. Les connexions sans flèches représentent des transitions bidirectionnelles

Les agronomes ont tout de suite adopté les diagrammes (cf. Fig. 5) qui visualisent les probabilités *a posteriori* des transitions entre états de Dirac qui représentent les probabilités de transition entre les OCS associées. Les probabilités de transitions entre états de Dirac e_i et e_j – associés aux OCS i et j – $\text{Prob}(Q_{t-1} = e_i, Q_t = e_j / O_1^T = o_1^T), 1 \leq t \leq T$ sont calculées lors de la dernière itération de l’algorithme Forward - Backward sur ce dernier type de HMM et sont représentées par des lignes brisées dès que leur probabilité atteint un certain seuil. L’épaisseur du trait est proportionnelle à la valeur de la probabilité. Une description complète de la construction des diagrammes – dits de Markov – se trouve dans EMS [LBBS⁺06].

2.5 Modèles pour représenter un champ spatial de Markov

Nous modélisons la mosaïque parcellaire par un champ de Markov dans lequel la parcelle porte une succession temporelle d’OCS. L’ensemble constitue une image de successions plutôt qu’une succession d’images et ne peut être étudié directement par les modèles numériques d’images appliqués au suivi de trajectoires d’objets [NCD07]. Dans le domaine spatial, la théorie des champs de Markov aléatoires – *Markov random field* (MRF) – propose de tenir compte des dépendances spatiales entre site en définissant des fonctions de potentiel qui quantifient l’attraction / répulsion des sites suivant leurs occupations. Les sites ayant les mêmes occupations s’attirent pour constituer un patch qui peut englober un site perdu au milieu possédant une autre occupation.

La représentation d’un paysage est donnée par un ensemble \mathcal{S} de sites et une relation de voisinage sur \mathcal{S} . Un site a 4 voisins – voisinage d’ordre 1 –

ou 8 dans le cas d'un système de voisinage d'ordre deux. $|\mathcal{S}|$ représente le nombre de sites et $\mathcal{N}(i)$ l'ensemble des voisins du site i . Nous appelons $\mathcal{E} = \{e_1, e_2, \dots, e_K\}$, l'ensemble des K différentes classes qui représenteront les "patches". $Z_i = e_k$ signifie que le site i est dans la classe e_k . L'ensemble des réalisations $\{Z_i = z_i\}$ s'appelle une *configuration*. Dans la suite, les variables aléatoires Z_i appartiennent à \mathcal{R}^K . En particulier, e_k est un vecteur binaire de \mathcal{R}^K ayant sa k^{eme} composante à 1, toutes les autres étant 0.

2.5.1 Le modèle de Potts avec champ externe

Dans un modèle de Potts avec champ externe, un unique paramètre $\beta > 0$ contrôle l'interaction – attraction *versus* répulsion – entre les classes alors qu'un vecteur supplémentaire V_i pondère l'importance donnée aux classes. La probabilité d'une configuration $Z = z$ est donnée par:

$$P(Z = z) = \frac{\exp\left(-\sum_{i \in \mathcal{S}} \left[z_i^t V_i - \beta \sum_{j \in \mathcal{N}(i)} z_i^t z_j\right]\right)}{W}.$$

W est un facteur de normalisation dont le calcul est impossible, d'où le besoin de recourir à des approximations telle celle du champ moyen. z^t représente la transposée de z et le produit $z_i^t z_j$ est égal à 1 si les sites i et j sont dans la même classe, 0 autrement.

2.5.2 L'approximation par champ moyen

Dans cette théorie, les probabilités d'appartenance d'un site à une classe se calculent en remplaçant les probabilités d'appartenance des sites voisins par leur valeur moyenne. Ainsi, Celeux [CFP03] propose de trouver la configuration optimale en résolvant l'équation de point fixe:

$$P_i^{mf}(e_s) = \frac{\exp\left[-V_i(e_s) + \beta \sum_{j \in \mathcal{N}(i)} P_j^{mf}(e_s)\right]}{\sum_{k=1}^K \exp\left[-V_i(e_k) + \beta \sum_{j \in \mathcal{N}(i)} P_j^{mf}(e_k)\right]}, \quad (4)$$

$V_i(e_k)$ étant la k^{eme} composante de V_i . Cette équation signifie que sur chaque site, la moyenne calculée par approximation du champ moyen est égale à la moyenne utilisée pour définir l'approximation.

Cette méthode de classification d'un territoire représenté par une matrice d'OCS est mise en œuvre dans le système ARPEnTage [MLB13]. Dans ce système, la probabilité $P_i^{mf}(e_s)$ qui quantifie l'appartenance du site i au patch e_s se calcule à l'aide de toute la séquence temporelle des OCS de ce site dont la probabilité est donnée par l'équation 3.

2.6 Prolongation d'une série temporelle par modèles de Markov

Le problème de la prolongation d'une série temporelle d'OCS peut être posé de la façon suivante : étant donné un territoire agricole représenté par un ensemble

de sites s_i possédant une suite d’occupations du sol en fonction du temps – soit $o_1^i, o_2^i, \dots, o_t^i$ la suite des occupations temporelles du site s_i – quelles seront les hypothèses vraisemblables pour o_{t+1}^i ? Nous nous plaçons dans un cadre probabiliste qui suppose que o_{t+1}^i suit la même loi de probabilité que o_t^i et que ses voisins spatiaux o_{t+1}^j (s_j est voisin de s_i). A l’intérieur de ce cadre, se pose le problème du choix de la période temporelle et de la portion d’espace pendant lesquelles sera estimée cette densité. La théorie des chaînes de Markov reste un formalisme élégant pour représenter et simuler des séquences. Dans ce travail, nous avons choisi de calculer la vraisemblance des séquences à l’aide de modèles de Markov cachés du second-ordre gauche-droite comme représentés figure 1 et mis en œuvre dans le logiciel CarottAge. Comme le représente la suite d’états associée à la suite d’observations, ces modèles permettent une segmentation temporelle du processus d’allocation des terres en le considérant stationnaire par morceaux (cf. Sec. 2.4). L’occupation au temps $t + 1$ est obtenue par un tirage au hasard en utilisant la densité $b_i(\cdot)$ du dernier état du HMM2 gauche-droite (cf. Fig. 1) qui modélise le dernier segment de la période d’étude.

Afin de tenir compte des phénomènes de rotations entre cultures choisies intentionnellement par les agriculteurs et pour pallier le fait que les HMM ne modélisent pas la corrélation entre observations successives, les observations sont des $n - \text{uplets}$ d’OCS (des triplés d’OCS relevés aux temps $t - 2$, $t - 1$ et t dans notre cas) qui se chevauchent d’année en année. Par exemple la séquence a, b, c, d, e, f, \dots est traitée comme : $abc, bcd, cde, def, \dots$. La Lorraine est connue pour être une région où règnent principalement des rotations de longueur 2 ou 3. La taille des $n - \text{uplets}$ détermine la quantité de mémoire pour représenter les densités $b_i(\cdot)$ donc le nombre de paramètres à estimer. Afin d’avoir un apprentissage robuste, nous nous sommes limités à $n = 3$.

La statistique sur le dernier état – soit N le dernier état – $b_N(o_{t-2}, o_{t-1}, o_t)$ est utilisée pour estimer la loi multinomiale $b(o_t/o_{t-2}, o_{t-1})$. Cette loi est utilisée pour générer o_{t+1} par tirage au sort en supposant connues les deux dernières observations o_{t-1} et o_t . Cette façon de faire correspond à la définition du modèle neutre de rotations tel que le définit Pontius [PHD04] au niveau de l’OCS et permet de simuler l’OCS à un pas de temps dans le futur.

3 Résultats

Différents modèles neutres ont été spécifiés et leurs capacités à prédire l’occupation d’un patch comparées. Ces modèles ont été calibrés sur différentes périodes d’étude et nous avons évalué leur performance pour prédire l’OCS l’année suivant la période d’étude. Parmi les différentes périodes d’étude utilisées se trouvent des périodes marquées par un changement dans la politique agricole menée par *Nestle Waters*. Le but de l’étude étant de corrélérer ces dates avec les ruptures détectées par les modèles neutres.

Dans un premier temps, à partir des données de 1996 à 2011, le système AR-PEnTAge a calculé une classification en 4 classes de patches notées e_2, e_3, e_4, e_5 – e_1 n’a pas d’existence agronomique et n’est utilisée que pour l’initialisation

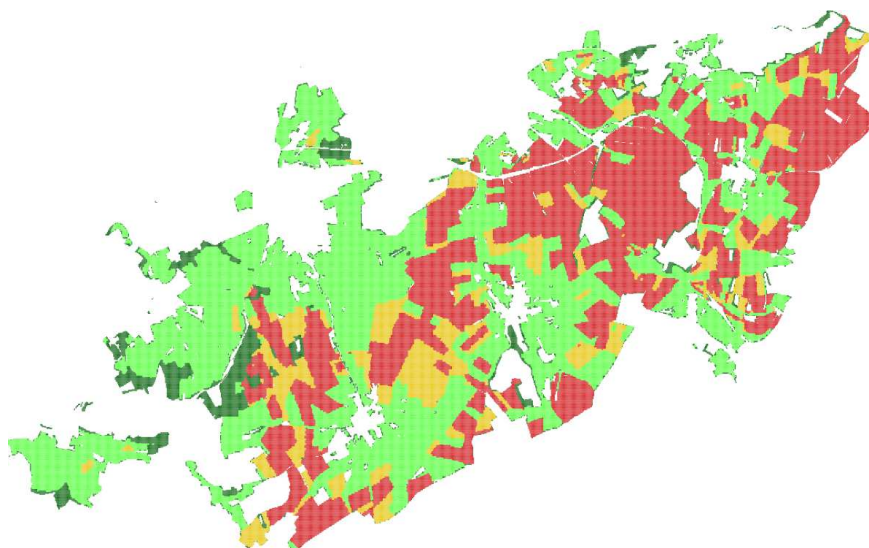


Figure 4: Le territoire de Vittel classé en 4 classes de patches calculées à partir des données de 1996 à 2010. Les patches verts clairs et verts sombres correspondent respectivement aux prairies et forêts pérennes. Les diagrammes de Markov des patches rouges et jaunes sont donnés dans les figures 5 et 6

des MRF— qui sont localisées (cf. Fig. 4) et qui permettent de séparer le corpus d’étude en autant de sous-corpus qui seront analysés indépendamment les uns des autres. Différents nombres de classes ont été essayés par l’analyste. L’expert n’en a retenu que quatre.

Les patches verts clairs et verts sombres correspondent respectivement aux prairies et forêts pérennes. Les deux autres classes, — dénotées par la suite *patch vert* et *patch rouge* — ont retenu l’attention de l’expert car étant porteuses d’enjeux environnementaux sur la qualité des eaux souterraines. Leurs diagrammes de Markov sont donnés Fig. 5 et Fig. 6

Dans les patches rouges, on retrouve des zones céréalières qui s’avèrent pour la plupart appartenir à des exploitations n’ayant pas encore contractualisé de nouveaux systèmes de culture plus prairiaux avec Agrivair. Dans cette classe de patches, le diagramme donné Fig. 5 montre entre les années 1996 et 2000 une diminution de la surface allouée aux prairies qui sont transformées en terre à blé. La mono culture du maïs se poursuit pendant la période 1996 – 2010 ainsi que la rotation maïs – blé. En 2005 la mono culture du colza disparaît et apparaissent la rotation colza – blé ainsi qu’une succession dans laquelle se retrouvent les prairies temporaires et le blé. Cette succession débute en fait dans les années 2000 car certaines prairies temporaires ont reçu l’étiquette luzerne par les enquêteurs entre 2003 et 2005 avant que cette dénomination ne soit abandonnée en 2006. Ces successions sont caractéristiques d’agriculteurs ayant contractualisé de nouveaux systèmes de culture plus prairiaux avec Agrivair.

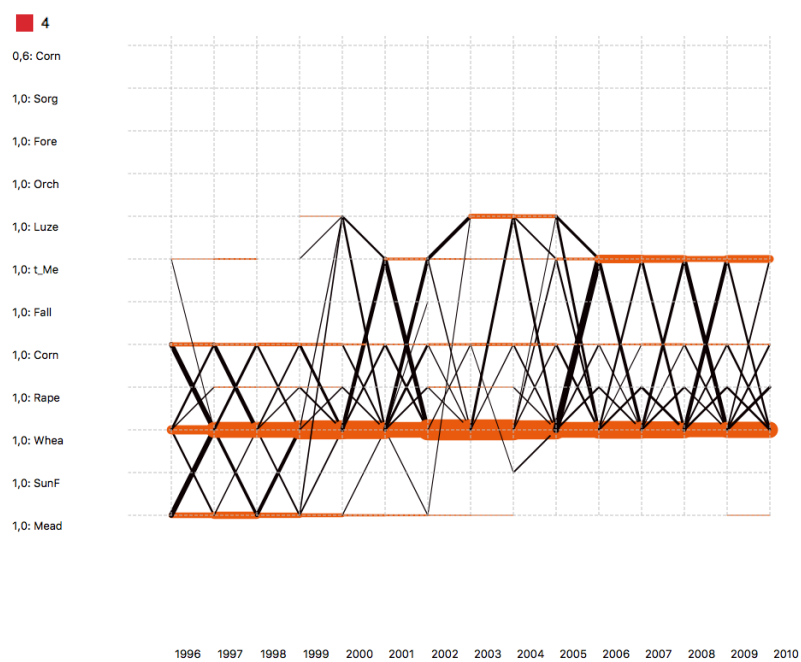


Figure 5: Diagramme de culture relatif au patch rouge (numéro 4). L'épaisseur du trait est proportionnelle à la probabilité de la succession

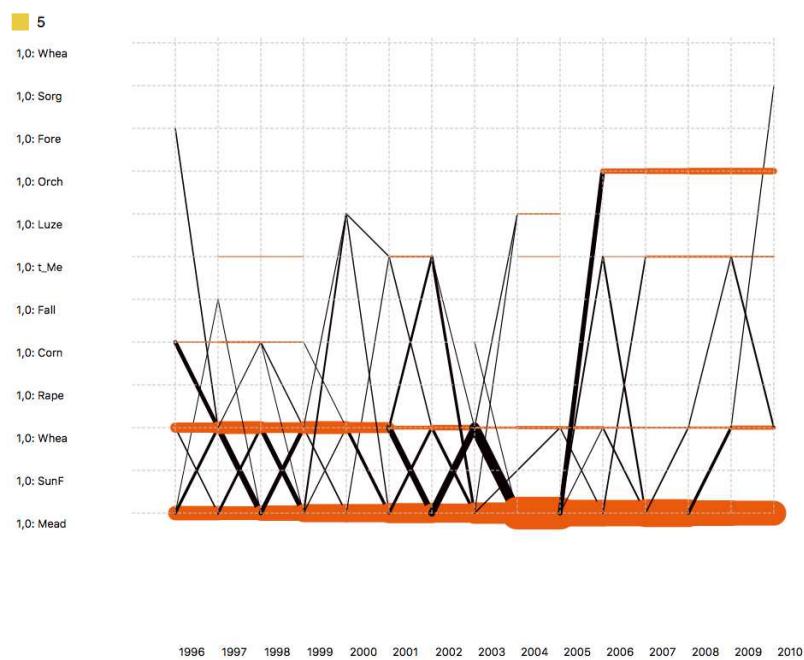


Figure 6: Diagramme de culture relatif au patch jaune (numéro 5)

Par opposition, la figure 6 associée aux patches jaunes montre une progression des prairies et une diminution des surfaces allouées au blé. Les transitions “Prairie temporaire – Prairie” et “Prairie” – “Blé” correspondent vraisemblablement à des confusions entre “prairie” et “prairie temporaire” dans les relevés d’enquêtes. On remarque la diminution de la mono culture du blé, la progression des prairies et l’absence de colza. L’année 2005 voit l’apparition de vergers correspondant à une nouvelle mesure environnementale proposée en contractualisation volontaire par *Nestle waters* (via Agrivair). La diminution au même moment de la surface des prairies est vraisemblablement due à l’apparition de ces vergers.

Sur chacun de ces deux sous corpus de données, nous avons spécifié différents modèles neutres sous forme de HMM2 linéaires (cf. Fig. 1) possédant un nombre croissant d’états : de 1 à 4. Comme dans ARPEntAge, nous utilisons des observations constituées par des triplés d’OCS se chevauchant. Un état de HMM2 est une distribution de triplés $P(O_{t-2}, O_{t-1}, O_t)$ qui permet de calculer la distribution conditionnelle $P(O_t/O_{t-2}, O_{t-1})$. o_{t+1} est généré par tirage aléatoire à partir de cette distribution multinomiale. Ces modèles ont été calibrés pour différentes périodes d’étude : de 1996 à 2010. Les apprentissages des HMM2 sont faits selon le principe du maximum de vraisemblance par le logiciel CarottAge. Ces HMM2 segmentent la période d’étude en autant de segments stationnaires qu’ils ont d’états et servent à simuler l’occupation des pixels l’année qui suit la fin de la période d’étude. Pour des raisons de robustesse de l’apprentissage, la première année testée pour la validation de la méthode est 2001 à partir des données de 1996 à 2000.

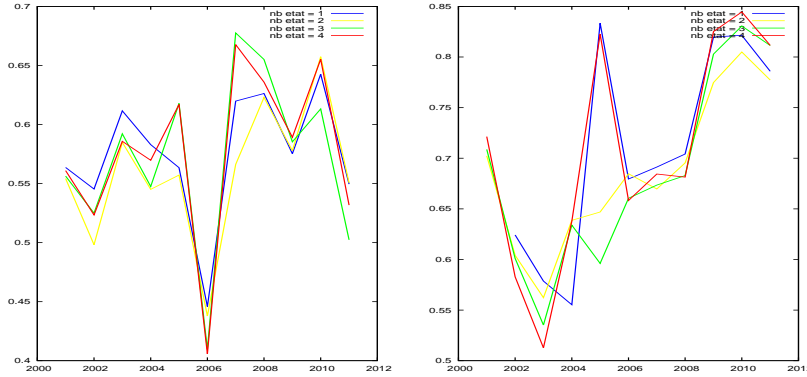


Figure 7: Taux de reconnaissance de pixels en fonction du temps sur les patches rouge et jaune. Le patch jaune contenant les exploitations conventionnées est plus prévisible que le patch rouge

La figure 7 montre les taux de classification des pixels simulés dans les patches rouges et jaunes pour les modèles neutres possédant différents nombres d’états. Ceux ci ont été calibrés sur les données de 1996 à t et évalués sur l’année $t + 1$.

On remarque que les taux sont faibles chaque fois qu’une discontinuité se

produit dans le processus d’allocation des terres, particulièrement en 2006 date à laquelle un verger de 300 ha a été installé.

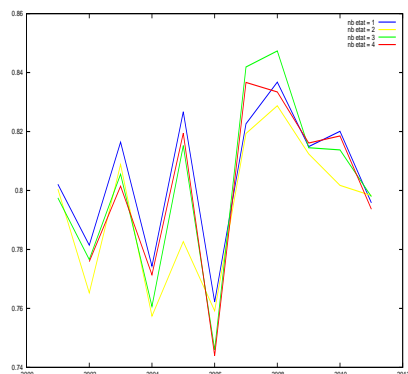


Figure 8: Par comparaison à la figure 7, taux de reconnaissance de pixels en fonction du temps sur la totalité du territoire

Par comparaison, la figure 8 montre le taux de classification sur tout le territoire sans utiliser l’information spatiale donnée par ARPEntAge. Les taux de classification sont bien meilleurs – environ 30 % supérieurs – à cause du biais introduit par les surfaces à couverture pérenne telles que forêts et prairies dont la prévision l’année suivante est immédiate. On remarque aussi que le modèle simple à un état a un comportement moyen honorable comparé aux modèles à plusieurs états. On voit que ce modèle s’adapte moins vite aux changements car ces derniers sont moyennés par des données passées devenues obsolètes. Sur ce modèle les ruptures dans le taux de reconnaissance sont différentes et moins bien visibles comparées aux modèles estimés sur les patch rouges et jaunes.

Les successions de culture des patches rouge et jaune ont des dynamiques différentes : 2006 est une rupture importante pour le patch rouge avec un pic en 2008 tandis que 2003 et 2010 sont les dates homologues pour le patch jaune. Le diagramme montrant le taux de reconnaissance sur tout le territoire ne permet pas cette analyse.

Les sous corpus utilisés pour calculer les taux de reconnaissance donnés dans la figure 7 utilisent donc l’information spatiale de 2011, ce qui constitue un biais dans la prédiction temporelle de l’année 2011. Pour éliminer ce biais, nous avons refait tourner ARPEntAge avec les données de 1996 à 2010 à la recherche de 4 classes. Le résultat est représenté dans la figure 4. La classification spatio temporelle reste similaire à celle obtenue en tenant compte de l’année 2011. Les patches occupent les mêmes places. Dans cette expérience, le découpage des parcelles en 2011 n’apparaît plus dans les données ce qui dégrade différemment les performances de prévision suivant le patch pour cette année. Le taux de reconnaissance du patch rouge reste inchangé à 54% tandis que le taux de reconnaissance du patch jaune perd 2% passant de 81% à 79%. Ces écarts sont significatifs au risque de 5% compte tenu de la taille des sous corpus.

	Mead	Wheat	Rapes	Corn	t_Me	Orch	Fore
Mead	0.84	0.04			0.12		
Wheat		0.76	0.01	0.08	0.15		
Rapes		0.87		0.07	0.05		
Corn	0.03	0.74		0.16	0.06		
t_Me		0.47		0.05	0.49		
Orch		0.01				0.96	
Fore						0.92	
Sorg		0.48			0.51		

Table 2: Matrice de confusion sur l'état rouge en 2011, $T_x = 0.54$

	Mead	Wheat	Rapes	Corn	t_Me	Orch	Fore	X
Mead	0.99							
Wheat	0.24	0.60		0.05	0.10	0.01		
Rapes	0.06	0.59		0.26	0.01	0.04		
Corn	0.05	0.76	0.01	0.11	0.07			
t_Me	0.25	0.39		0.03	0.34			
Orch						1,00		
Fore							0.22	0.78
Sorg	0.01	0.94			0.05			

Table 3: Matrice de confusion sur l'état jaune en 2011, $T_x=0.72$

Les matrices de confusion données Tab. 3 et Tab. 3 sont calculées pour un HMM2 à 4 états calibré avec les données temporelles et spatiales de 1996 à 2010 et testé sur l'année 2011.

La matrice associée à l'état rouge montre les principales confusions suivantes. (i) Là où on attendait un colza, c'est un blé qui a été observé. Vraisemblablement parce que le colza est cultivé de moins en moins en 2001. (ii) Là où on attendait une prairie temporaire, on a observé un blé ce qui s'explique par le fait qu'il n'y a pas d'alternance régulière entre ces OCS.

La matrice de confusion de l'état jaune donnée Tab. 3 montre que ce patch dans lequel se situe un grand nombre d'exploitations sous contrat est mieux prédit que le patch rouge. La grande proportion de forêts (0.78 %) étiquetée comme "Non Classée" (=X) provient de sites situés en périphérie du territoire et qui n'appartenaient plus au territoire en 2011. La même grille *raster* d'échantillonnage n'a pas pu être utilisée entre 1996 et 2011. Le travail d'enquête est un labeur de longue haleine plus ou moins bien documenté et a été effectué par un grand nombre de personnes. Le recalage des SIG présente quelques approximations.

La figure 9 localise les erreurs de prédiction de 2011. Les différentes teintes bleues représentent les OCS erronées prédites. Les patches qui ont gardé leur couleur d'origine sont constitués de parcelles correctement prédites. L'apprentissage des modèles stochastique (HMM2 et MRF) été fait sur les données de la période 1996 – 2010. L'analyse visuelle des patches fait penser qu'un phénomène de redécoupage de parcelles a eu lieu en 2011. Le patch erroné est quelque fois à

	Mead	Wheat	Rape	Corn	t_Me	Orch	Fore
Mead	1.00						
Wheat	0.05	0.78		0.05	0.12		
Rape	0.02	0.88		0.06	0.04		
Corn	0.06	0.71		0.16	0.06		
t_Me	0.03	0.48		0.04	0.44		
Orch						0.99	
Fore							0.93
Sorg	0.01	0.52			0.44		0.02

Table 4: Matrice de confusion sur tout le territoire en 2011, Tx=0.79

l'intérieur d'une zone correctement prédite et ses frontières sont linéaires. Ceci est dû en partie au fait que l'analyse spatiale effectuée par ARPEnTage n'a pas utilisé les données de l'année 2011 et que les **MRF** ne font aucune prédiction sur les frontières de parcelles. Le processus de redécoupage n'est pas modélisé.

4 Discussion

La modélisation markovienne des processus siégeant à différentes échelles dans la mosaïque des parcelles agricoles a été souvent adoptée [CWB⁺08, CMA⁺10, LMB09] et critiquée. En effet, Gaucherel [GBH⁺12] soutient que le paysage est un objet temporel non markovien. Dans notre travail, nous l'avons partiellement suivi en montrant que la stationnarité par morceaux des successions de cultures révélait des changements non markoviens de politiques menées par Agrivair. Les modèles neutres utilisant le principe de localité mis en œuvre dans des **HMM2** s'avèrent être des bons outils pour tester cette hypothèse de localité et continuer à s'ajuster à des données non markoviennes.

Comparés à des **HMM1**, les **HMM2** modélisent mieux les phénomènes de durée sans avoir la complexité des **HMM** semi continus [Lev86, Gué05, CP08]. Les **HMM** généralisent les matrices de Markov et possèdent des algorithmes d'apprentissage automatique qui suppriment le choix de la période de calibration. En revanche, quand le nombre d'états ne correspond pas au nombre de périodes stationnaires, les paramètres des modèles donnent des prédictions bien moins bonnes comparées à celles d'un modèle plus simple à un état.

La modélisation des interactions spatiales peut se faire aussi à l'aide d'automates cellulaires [HHM06, BOC⁺14, Kar05]. Les cellules permettent de prendre en compte de nombreuses contraintes par opposition au **MRF** dans lesquels seulement une fonction potentiel peut être définie [GFA⁺06].

La communauté scientifique a déjà proposé différents logiciels pour évaluer les changements d'usage des sols. Elias [EDB12] propose une synthèse des travaux de modélisation des changements d'usage des terres dans les pays en voie de développement. Mas [MKP⁺14] a comparé quatre produits logiciels : **CLUE-S**, **DINAMICA EGO**, **CA_MARKOV** et **LCM** sur un terrain d'étude virtuel dans le but de construire des cartes de changements potentiel d'OCS. Une poursuite de

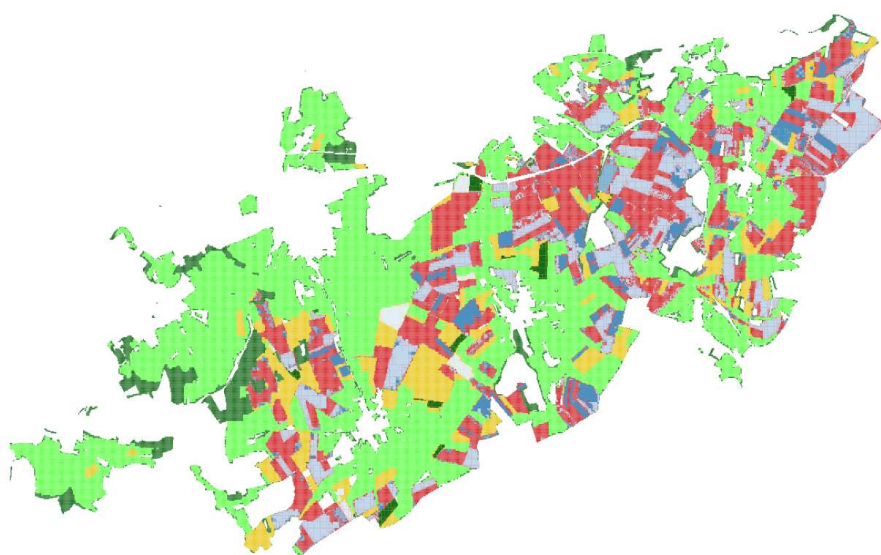


Figure 9: Carte obtenue par simulation prospective à partir des données de 1996 à 2010. Les erreurs de prédiction en 2011 sont représentées en bleu. Les patches qui ont gardé leur couleur d'origine sont constitués de parcelles correctement prédites

notre travail serait d’ajouter ARPEnTAge et CarottAge à la liste des logiciels donnés plus haut pour permettre leur usage à une communauté plus large.

Dans le domaine de la simulation inductive de structures paysagères, l’unité allouée est la majeure partie des cas une OCS dont la présence / absence est projetée. Par exemple, Pontius [PHD04] s’intéresse à l’agrandissement de l’emprise des villes dans les campagnes ou au phénomène de déforestation en attribuant deux modalités à un site : “Change” et “No change”. Notre analyse s’intéresse aux changements dans les dynamiques et traite des successions d’OCS quelquefois structurée en rotations telles que “Colza - blé” ou “prairie temporaire - prairie temporaire - blé”. Par exemple, la poursuite au temps $t + 1$ d’une rotation triennale commencée en $t - 1$ correspond toujours à l’utilisation d’un modèle “neutre” malgré le changement d’OCS. Des résultats plus complets pourraient être obtenus en utilisant des quadruplets d’OCS à la place de triplets.

L’évaluation des résultats d’une simulation prospective dans une mosaïque agricole a fait l’objet de nombreux travaux [JS01, PHD04, BCMG15]. Nous nous sommes limités au taux de correcte prédiction du pixel donné dans les tableaux 3, 3 e 3. cohérente aurait été de compter les parcelles correctement prédites. Mais celles ci changent de forme dans le temps comme le laisse supposer la figure 9. Nos essais de définition de micro-parcelles – portions de parcelles qui portent la même succession pendant la durée de l’étude – agencées dans un système de voisinage irrégulier [MLM10] n’ont pas donné de résultats concluants. Le krigeage des parcelles faisait perdre l’information de superficie très importante en agriculture. Les taux donnés dans cet article se rapportent à des superficies correctement prédites. Le problème du re-découpage d’une parcelle, actuellement traité comme une erreur de prédiction, devrait être pris en compte par un modèle neutre car il ne révèle pas une instabilité dans la mosaïque agricole et constitue une pratique courante pour adapter un rythme de production à une surface agricole. Dans cet ordre d’idée, il faut citer les travaux de planification de Akplogan [AGS⁺13] qui cherche à placer un ensemble de cultures à l’échelle de la parcelle, l’ilot et l’exploitation.

En Écologie il existe une controverse au sujet des définitions et des différences entre modèle nul et modèle neutre [Pig04, GM06] qui, suivant le cas, modélisent la stabilité des processus sur un territoire ou des comportements d’individus à l’intérieur d’un territoire. Nous avons adopté le terme de modèle neutre dans la lignée des travaux de Gaucherel [LBLA⁺09] et Le Ber [LBLA⁺09].

5 Conclusions

Nous avons présenté une méthode générique pour simuler les dynamiques d’occupation du sol qui s’appuie sur des modèles stochastiques. Le même formalisme – chaînes et champs de Markov – permet de modéliser l’évolution temporelle et spatiale. Le territoire est représenté comme un assemblage de “patches” qui sont eux-mêmes représentés par des images “rasterisées”. Nous simulons l’occupation d’un pixel sur la base des occupations des 2 années précédentes et des pixels voisins. La simulation se fait par tirage aléatoire des occupations à l’aide de den-

sités estimées sur une région spatiale et une période temporelle choisies de façon optimale selon le maximum de vraisemblance. Le résultat de la simulation peut être vu comme la constitution d’un territoire “neutre” dans lequel se poursuit la tendance actuelle. Dans la simulation d’un patch homogène de successions de culture, les HMM2 avec un nombre d’états qui correspond au nombre de périodes stationnaires donnent de meilleurs résultats qu’un modèle stationnaire simple à un seul état.

La méthode proposée permet d’identifier des changements non pris en charge par le modèle (cf. Fig. 9). Les écarts au modèle représentés par les courbes de la figure 7 montrent au cours du temps des variations. L’écart entre le modèle “neutre” et la réalité permet de déceler des ruptures dans le processus d’allocation des terres par les agriculteurs. Sur ces territoires, la variable explicative la plus plausible serait la dynamique des signatures des contrats par les agriculteurs impliquant un cahier des charges dans la conduite des cultures. Ce modèle permet non seulement d’évaluer la qualité prédictive de la simulation, mais surtout de spatialiser cette prédiction. Ainsi, il est possible de caractériser des facettes de paysage suivant leur sensibilité à cette prédiction. Ce fait nous permet d’accéder aux patchs où les simulations sont les plus fiables ... ou les moins fiables. L’hétérogénéité d’un paysage peut ainsi se lire sur ce critère.

Nous remercions Agrivair pour son aide à la captation des informations et au programme de recherche AGREV 3 ainsi que les agriculteurs de l’impluvium pour leur disponibilité

References

- [AGG⁺00] C. Agarwal, G. L. Green, M. Grove, T. Evans, and C. Schweik. A review and assessment of land-use change models: dynamics of space, time and human choice. In *Ecosystems And The North American Carbon Cycle 9 Human Choice. Cipec Collaborative Report Series No. 1, Center For The Study Of Institutions, Populations, And Environmental Change, Indiana University And The Usda Forest Service*, pages 10–1111, 2000.
- [AGS⁺13] M. Akplogan, De Givry, Metivier S., Quesnel J.-P., A. G., Joannon, and F. Garcia. Solving the crop allocation problem using hard and soft constraints. *RAIRO Recherche Opérationnelle*, 47(2):151–172, 2013.
- [BCMG15] Vincent Bonhomme, Mathieu Castets, Jules Morel, and Cédric Gaucherel. Introducing the vectorial kappa: An index to quantify congruence between vectorial mosaics. *Ecological Indicators*, 57:96 – 99, 2015.
- [BOC⁺14] Reine Maria Basse, Hichem Omrani, Omar Charif, Philippe Gerber, and Katalin Bódis. Land use changes modelling using advanced methods: Cellular automata and artificial neural networks.

- the spatial and explicit representation of land cover dynamics at the cross-border region scale. *Applied Geography*, 53:160 – 171, 2014.
- [BRM⁺12] Marc Benoît, Davide Rizzo, Elisa Marraccini, Anna Camilla Moonen, Mariassunta Galli, Sylvie Lardon, Hélène Rapey, Claudine Thenail, and Enrico Bonari. Landscape agronomy: a new field for addressing agricultural landscape dynamics. *Landscape Ecology*, 27(10):1385–1394, 2012.
- [CFP03] Gilles Celeux, Florence Forbes, and Nathalie Peyrard. EM procedures using mean field-like approximations for Markov model-based image segmentation. *Pattern Recognition*, 36(1):131–144, 2003.
- [CMA⁺10] M.S. Castellazzi, J. Matthews, F. Angevin, C. Sausse, G.A. Wood, P.J. Burgess, I. Brown, K.F. Conrad, and J.N. Perry. Simulation scenarios of spatio-temporal arrangement of crops at the landscape scale. *Environmental Modelling & Software*, 25(12):1881 – 1889, 2010.
- [CP08] Florence Chaubert-Pereira. *Markov and semi-Markov switching regression models. Application to forest tree growth*. Theses, Université Montpellier II - Sciences et Techniques du Languedoc, November 2008.
- [CWB⁺08] M. Castellazzi, G. Wood, P. Burgess, J. Morris, K.F. Conrad, and J.N. Perry. A systematic representation of crop rotations. *Agricultural Systems*, 97:26–33, 2008.
- [DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum-Likelihood From Incomplete Data Via The EM Algorithm. *Journal of Royal Statistic Society, B (methodological)*, 39:1 – 38, 1977.
- [EDB12] Peter Elias, Samuel Dekolo, and Olatunji Babatola. Land Use Change Modelling In Developing Countries: Issues And Prospects. *International Journal of Geography and Geology*, 1(1):23–41, 10-2012 2012.
- [For95] Richard T.T. Forman. Some general principles of landscape and regional ecology. *Landscape Ecology*, 10(3):133–142, 1995.
- [GBH⁺12] Cédric Gaucherel, Frédéric Boudon, Thomas Houet, Mathieu Castets, and Christophe Godin. Understanding patchy landscape dynamics: Towards a landscape language. *Plos One*, (9), 2012. DOI : 10.1371/journal.pone.0046064.
- [GFA⁺06] C. Gaucherel, D. Fleury, D. Auclair, P. Dreyfus, and Gaucherel Cédric. Neutral models for patchy landscapes. *Ecological Modelling*, 197(1):159–170, August 2006.

- [GM06] N. J. Gotelli and Brian J. McGill. Null versus neutral models: What’s the difference? *Ecography*, 29(5):793–800, 2006.
- [Gué05] Yann Guédon. Hidden hybrid markov/semi-markov chains. *Computational Statistics and Data Analysis*, 49(3):663 – 688, 2005.
- [HHM06] Thomas Houet and Laurence Hurbert-Moy. Modelling and projecting land-use and land-cover changes with a cellular automaton in considering landscape trajectories. *EARSeL eProceedings*, 5(6):63–76, 2006.
- [Hou15] Thomas Houet. Usages de modèles spatiaux pour la prospective. *Revue internationale de Géomatique*, 25(1):123 – 143, 2015.
- [Hub01] S.P. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Monographs in Population Biology. Princeton University Press, 2001.
- [JS01] R.Gil Pontius Jr. and Laura C Schneider. Land-cover change model validation by an {ROC} method for the ipswich watershed, massachusetts, {USA}. *Agriculture, Ecosystems & Environment*, 85(1–3):239 – 248, 2001. Predicting Land-Use Change.
- [JSMP06] A. Joannon, V. Souchère, P. Martin, and F. Papy. Reducing runoff by managing crop location at the catchment level: considering agro-economic constraints at farm level. *Land Degradation and Development*, 17(5):467–478, 2006.
- [Kar05] Jarkko Kari. Theory of cellular automata: A survey. *Theoretical Computer Science*, 334(1):3 – 33, 2005.
- [LBBS⁺06] F. Le Ber, M. Benoît, C. Schott, J.-F. Mari, and C. Mignolet. Studying Crop Sequences With CarrotAge, a HMM-Based Data Mining Software. *Ecological Modelling*, 191(1):170 – 185, Jan 2006.
- [LBLA⁺09] F. Le Ber, C. Lavigne, K. Adamczyk, F. Angevin, N. Colbach, J.-F. Mari, and H. Monod. Neutral modelling of agricultural landscapes by tessellation methods – application for gene flow simulation. *Ecological Modelling*, 220(24):3536 – 3545, December 2009. <http://hal.archives-ouvertes.fr/hal-00409081/fr/>.
- [Lev86] S. E. Levinson. Continuously Variable Duration Hidden Markov Models for Automatic Speech Recognition. *Computer Speech and Language*, 1:29 – 45, 1986.
- [LMB09] E.G. Lazrak, J.-F. Mari, and M. Benoît. Landscape regularity modelling for environmental challenges in agriculture. *Landscape Ecology*, 25(2):169 – 183, Sept. 2009.

- [MHK97] J.-F. Mari, J.-P. Haton, and A. Kriouile. Automatic Word Recognition Based on Second-Order Hidden Markov Models. *IEEE Transactions on Speech and Audio Processing*, 5:22 – 25, January 1997.
- [MKP⁺14] Jean-François Mas, Melanie Kolb, Martin Paegelow, María Teresa Camacho Olmedo, and Thomas Houet. Inductive pattern-based land use/cover change models: A comparison of four software packages. *Environmental Modelling & Software*, 51:94 – 111, 2014.
- [MLB13] Jean-François Mari, El-Ghali Lazrak, and Marc Benoît. Time space stochastic modelling of agricultural landscapes for environmental issues. *Environmental modelling & software*, 46:219–227, August 2013. http://hal.inria.fr/hal-00807178/PDF/arpentage_hal.pdf.
- [MLM10] Jean-Francois Mari, El Ghali Lazrak, and Benoît Marc. Fouille de paysages agricoles: analyse des voisinages des successions d’occupation du sol. In Gérard Ligozat et Odile Papini Maroua Bouzid, Florence Le Ber, editor, *Colloque RTE (Raisonnement sur le Temps et l’Espace) en marge de RFIA 2010*, Caen France, 01 2010. http://hal.archives-ouvertes.fr/inria-00482811/PDF/RTE10-Mari_et_al-13Janv.pdf.
- [MSB07] C. Mignolet, C. Schott, and M. Benoît. Spatial dynamics of farming practices in the Seine basin: Methods for agronomic approaches on a regional scale. *Science of The Total Environment*, 375(1–3):13–32, April 2007.
- [NCD07] V. Noyon, C. Claramunt, and D.Devogele. A relative representation of trajectories in geographical spaces. *Geoinformatica, springer*, 14(11):479 – 496, 2007.
- [PHD04] R. G. Pontius, Diana Huffaker, and Kevin Denman. Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling*, 179(4):445 – 461, 2004.
- [Pig04] Simone Pigolotti. *Neutral theories in ecology: species abundance and extinction dynamics*. PhD thesis, Scuola Internazionale Superiore di Studi Avanzati - Trieste ITALY, 2004. http://www.nbi.dk/pigo/articles/Pigo_Phddthesis.pdf.
- [RJ95] L. Rabiner and B. Juang. *Fundamentals of Speech Recognition*. Prentice Hall, 1995.
- [SLM⁺12] Noémie Schaller, El Lazrak, Philippe Martin, Jean-François Mari, Christine Aubry, and Marc Benoît. Combining farmers’ decision rules and landscape stochastic regularities for landscape modelling. *Landscape Ecology*, 27:433–446, 2012.

- [SLOW11] Adrian Southern, Andrew Lovett, Tim O’Riordan, and Andrew Watkinson. Sustainable landscape governance: Lessons from a catchment based study in whole landscape design. *Landscape and Urban Planning*, 101(2):179 – 189, 2011.
- [VL01] A Veldkamp and E.F Lambin. Predicting land-use change. *Agriculture, Ecosystems & Environment*, 85(1–3):1 – 6, 2001. Predicting Land-Use Change.